

機械翻訳(MT)を学習ツールとするための考察

— 英作文を学ぶ大学生を対象とした量的・質的調査 —

研究者:湯浅 麻里子 大阪府／関西大学大学院 在籍

《研究助言者:寺内 一》

概要

本調査は、習熟度が低い大学生英語学習者がどのように機械翻訳(MT)を使っているのか、現状をマクロとミクロの視点で明らかにした。まず、773名の参加者による質問紙データを用いて、MTを使ってライティングを行う際のエンゲージメントを包括的に測る「MTエンゲージメント」尺度を作成した。18項目から成る尺度は、行動、プリエディット認知、ポストエディット認知、情意、社会の5つの因子で構成され、信頼性と妥当性が確認された。続いて尺度の5つの変数に基づき、434名の参加者を類似するグループに分けた。参加者は、①低エンゲージメント、②中エンゲージメント、③高エンゲージメントの3つのグループに分かれた。MTを使って英文を書くことに対する自己効力感は、高エンゲージメント・グループが高く、中・低エンゲージメント・グループが低かった。さらにインタビューから、7名の学生がそれぞれの文脈の中で工夫しながらMTを使っていたことがわかった。

1 はじめに

機械翻訳(MT)は2016年、AIを用いたニューラル機械翻訳(NMT)の導入を機に、翻訳精度が劇的に向上した(Ducar & Schocket, 2018; Stapleton & Kin, 2019)。MTは私たちの日常に浸透し、グローバル社会の中で言語の壁を打ち破っている。その影響は教育現場も例外ではない。特に大学生の外国語(L2)ライティングにおけるMT使用は世界的に広がり(Briggs, 2018; O'Neill, 2019), “MT-enabled L2 writing” (MTを活用したL2ライティング; Jiang et al., 2024, p.1) が標準化しつつある。MTを取り入れることで、L2ライティングの質が高まり(Chon et al., 2021; Lee, 2020; Tsai, 2019), 語彙の使用と習得にも役立つことが報告されている(Lo, 2023)。これらの研究の多くは、ヨーロッパ言語共通参照枠(Common European Framework of Reference for Languages; CEFR)のB1以上の中・上級学習者を対象としており、MTは習熟度が高い学習者に効果があるとされている(Klimova et al., 2022)。一方、日本では大学生のおよそ8割がCEFR A2レベルで(Negishi et al., 2013), このような初級学習者を対象とした研究は世界でも限られている(Lee, 2021)。日本人大学生の多くが英語の課題にMTを使っているようだが(Yamada et al., 2021), 授業での活用例はまだ少ない。しかし、AI技術の急成長と実社会での役割を考えると、大学生のMT使用はもはや見過ごせない(Godwin-Jones, 2022; Kern, 2024)。

MTをL2学習に役立てるためには、入力時に「MTが出力しやすい母語(L1)」(“translation-friendly writing”)を考え、L1を修正するプリエディット(前編集)や、MTが出力したL2の文法や語彙の誤りを修正するポストエディット(後編集)などの方略を使うことが重要である(Bowker & Buitrago-Ciro, 2019)。つまり、ライティングのプロセスにどれだけ深く考えながらたずさわるか、「エンゲージメント」

が学習の鍵となる。しかし、日本の教育現場ではMT使用の事例も乏しく、学生のMTの使い方はあまり理解されていない(Gally, 2019)。そこで本調査は、エンゲージメントの枠組みを使い、大学生がどのようにMTを使っているのか、現状をマクロとミクロの視点から明らかにする。エンゲージメントとは、学習プロセスへ積極的に参加することで、行動、認知、情意、社会などの複数の要素で構成される(Hiver et al., 2021)。エンゲージメントはL2学習を促進することが報告されている(Lei et al., 2018; Liu et al., 2023)。大学生のMT使用の現状をエンゲージメントの観点から包括的に捉えることで、教師は学生のニーズや課題に気づき、教授法を見直すきっかけになるのではないかと。本調査では、非英語専攻の大学生の多くに相当するCEFR A2レベルの学習者を対象に、MTを使ってライティングを行うことへのエンゲージメント(「MTエンゲージメント」と略す)を測る尺度を作り、大学生の現状を数値で概観する。さらに個人に焦点を当ててインタビューを行い、学習者がそれぞれの文脈の中でどのようにMTを使っているのかを、より具体的に描写することを試みる。

2 先行研究

2.1 外国語学習とエンゲージメント

教育心理学から発展したエンゲージメント(Reschly & Christenson, 2022)は、近年、第二言語習得の分野で動機づけと学習行動を結びつける概念として注目を集めている(Hiver et al., 2021; Mercer & Dörnyei, 2020)。エンゲージメントは主に行動(努力や時間)、認知(深い思考や方略使用)、情意(学習に対する態度や感情)の3つの要素から構成され(Mercer, 2019)、学習者の行動、認知、情意がどのように促されて維持されるかを表す(Schunk & DiBenedetto, 2022)。さらに、社会(L2での他者とのやりとり; Svalberg, 2009)や主体性(学習への積極的な関わり; Reeve & Tseng, 2011)が構成要素に含められることもある。外国語を習得するには、語彙や文法を理解するだけでなく、持続的にL2を使用することが不可欠である(Hiver et al., 2024)。このように、エンゲージメントは学習行動を多面的な観点で捉えることができる。エンゲージメントとL2学習の成果には相関があることも実証されている(Lei et al., 2018; Liu et al., 2023)。

2.2 機械翻訳(MT)とエンゲージメント

習熟度が高い英語学習者がMT訳を参照しながら自作の英文を修正すると、語彙や文法の誤りが減り、高度な語彙や複雑な構文が増えることが報告されている(Chon et al., 2021; Lee, 2020; Tsai, 2019)。しかしMTは、L1を入力するだけで瞬時に英文ができる。MTを使うときは、ポストエディットを行うことが推奨されるが(Bowker & Buitrago-Ciro, 2019)、習熟度の低い学生はMT出力の誤りを見つけることが自体が困難である(Kol et al., 2018; Shin & Chon, 2023)。また、ポストエディット方略の使用頻度も習熟度に影響される(Chung, 2020; Shin & Chon, 2023)。日本人大学生のおよそ8割を占めるCEFR A2レベル(Negishi et al., 2013)の英語学習者によるMT使用は、学習へ悪影響を及ぼすと懸念されている。

コンピュータの登場以来、L2学習はテクノロジーの恩恵を受けてきた。私たちの日常に存在しているテクノロジーを排除することは学習環境を制限してしまう(Chun et al., 2016)。MTも例外ではない。L2学習にMTを取り入れることはごく自然ではないか。MTは学習者を惹きつけ、エンゲージメントを高めるタスクを作ることができる(Kern, 2024)。学習者は有意義で価値あるタスクには意欲的に取り組む(Williams & Burden, 1997)。このような学習機会を作るためには、教師は大学生がどのようにMTを使ってライティングをしているのか、現状を知ることが不可欠である。エンゲージメントの枠組みを使うことで、学生のMT使用の実態を行動、認知、情意、社会などのさまざまな観点から包括的に理解することができるだろう。

2.3 エンゲージメントと自己効力感

学習に大きな影響を与える要因の1つである自己効力感とは、エンゲージメントと強く関係している。自己効力感とは、ある課題に対する自身の能力へのビリーフで、自分はできると信じることである(Bandura, 1997)。このビリーフが課題の選択や努力に影響を与え、学習成果を高める(Pajares, 1996)。教育心理学では、自己効力感と動機づけ、エンゲージメント、学習成果との強い相関が実証されている(Schunk & DiBenedetto, 2022; Zimmerman, 1995)。

教育心理学における自己効力感の知見は、第二言語習得においても広く研究されており、自己効力感と外国語4技能の習得には強い相関があることがわかっている(Goetze & Driver, 2022; Wang & Sun, 2020)。自己効力感は単独で機能するのではなく(Graham, 2022)、自己調整学習や学習方略などと連動して機能する(Bai & Wang, 2023; Sun & Wang, 2020)。リーディングやライティングなどの方略指導が自己効力感を高めることも実証されている(Graham et al., 2020; Teng & Zhang, 2020)。エンゲージメントと自己効力感においても正の相関がある。Cai and Xing(2023)は、L2学習に対する自己効力感が、認知エンゲージメントを通じて成果を高めることを示している。Tsao(2021)は、L2ライティングに対する自己効力感が、フィードバックに対するエンゲージメントを高めることを報告している。自己効力感がエンゲージメントを高め、パフォーマンスが向上し、達成感や自己効力感を促すことで、さらなる学習の好循環が生み出されるのである。

2.4 調査の目的と構成

本研究は、CEFR A2レベルに相当する非英語専攻の大学生英語学習者が、MTを使ってどのようにL2ライティングを行っているのかを、マクロとミクロの視点から調査する。調査1では、CEFR A2レベルのL2学習者に特化した「MTエンゲージメント」尺度を作成する。調査2では、大学生のMT使用の実態を調査1で作成した尺度による量的手法と、インタビューによる質的手法により明らかにする。

3 調査1: MTエンゲージメント尺度の作成

3.1 目的

調査1では「MTエンゲージメント」尺度を作成する。エンゲージメントは個々の学習者が置かれた文脈に依存する(Hiver et al., 2024)。妥当性の高い尺度を開発するため、2つのリサーチクエスション(RQ)を設定した。

RQ1 MTエンゲージメント尺度はどのような因子で構成されるか。

RQ2 MTエンゲージメント尺度の信頼性と妥当性はどの程度か。

3.2 方法

3.2.1 参加者

調査者が所属する大学の17学部773名の学生(男性387名、女性372名、回答しない14名)が本調査への協力に同意し、参加した。参加者は非英語専攻の学生を対象とした必修英語クラスの1年生(392名)と2年生(381名)で、英語習熟度の判定には、GTEC-LRによるプレスメントテスト(500点満点)の結果($N = 717$,

$M = 220.65$, $SD = 46.91$)を利用した。点数の分布からおよそ9割がCEFR A2レベルを占め、少数のA1とB1レベルを含んでいた。参加者は平均的な非英語専攻の大学生英語学習者を表していると言える。

データ収集に際しては、回答は無記名で、データはすべて統計的に処理し、個人が特定されないことを説明した。調査への協力に同意した学生はスマートフォンを使ってGoogle Formsから回答した。データ収集に先立ち、調査者が所属する大学院の倫理委員会に許可を得た。

3.2.2 測定道具

本調査では2つの質問紙を用いた。本調査のために作成した「MTエンゲージメント」尺度と「MTを用いたL2ライティングに対する自己効力感(「MT自己効力感」と略す)」尺度(Yuasa & Takeuchi, 2022)である。エンゲージメントが自己効力感とどの程度の相関があるか(併存的妥当性)を調べるためである。

3.2.3 実施手順

MTエンゲージメント尺度の作成は、項目の作成、尺度の作成、尺度の検証の3段階で行った(Boateng et al., 2018)。尺度の作成に探索的因子分析を、尺度の検証に検証的因子分析を実施した。第1段階では、MTエンゲージメントの5つの下位概念を定義した。「MTを使う時間と努力」を表す行動、「MT出力を向上させる方略使用」を表す認知、「MT使用への態度や感情」を表す情意、「学習者間のMTについてのやりとり」を表す社会、「積極的なMT使用と個人化」を表す主体性である。これらの定義と質的データによる学習者の方略使用に基づき(Yuasa & Takeuchi, in press), 30項目の草案を作成した。これらの項目は、第二言語習得分野の研究者と議論を通して修正を重ね、内容的妥当性を高めた。

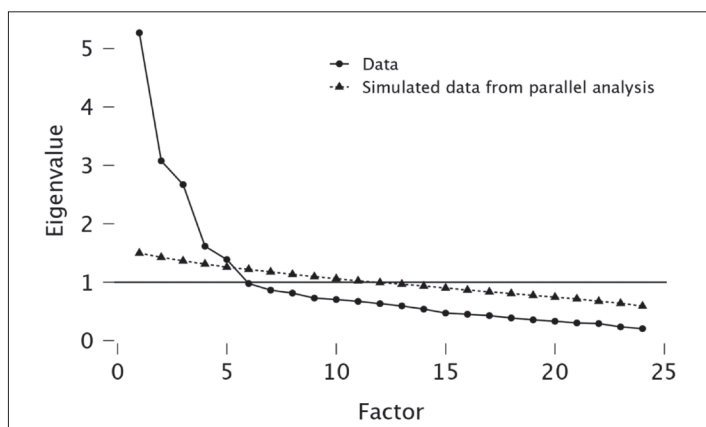
第2段階は尺度の作成である。2023年6月に108名、7月に269名の学生を対象に予備調査を実施した。その結果を受けて項目の改善を行い、24項目の質問紙を作成した。この質問紙を同年9月から10月に773名の学生に実施した。「あなたは普段、英文を書くときにどのように機械翻訳を使っていますか」という教示文に対して5件法(1. 当てはまらない 2. あまり当てはまらない 3. どちらとも言えない 4. やや当てはまる 5. 当てはまる)での回答を依頼した。データクリーニングにより、テストスコアの欠如や不自然な回答(すべての項目が同じ回答)、外れ値などを除いた708名のサンプルを、探索的因子分析と検証的因子分析のために、354名ずつの2グループに分けた。各グループ100名ずつを無作為抽出して t 検定を行い、習熟度に差がないことを確認した($t = .55$, $df = 198$, $p = 0.59$)。354名のデータを基に、因子分析の前提条件である、サンプル数が十分であること(KMOの標本妥当性指標 = .82)と、変数間に相関があること(バートレットの球面性検定: $\chi^2 = 3238.71$, $df = 276$, $p < .001$)を確かめた後、JASP(ver.0.18)を用いて探索的因子分析を実施した。因子負荷量の最低値はHair et al.(2019)に基づき.50とした。

第3段階では、探索的因子分析で特定した因子構造の内的―貫性と尺度の妥当性を、検証的因子分析により検証した。尺度項目の内的―貫性にはクロンバック α の信頼性係数を、また、妥当性は構成概念妥当性、収束的妥当性、弁別妥当性を調べた。さらにMT自己効力感尺度を用いて、外的な基準との間にどの程度相関があるのかを検証する併存的妥当性を調べた。

3.3 結果

3.3.1 MTエンゲージメントの因子構造(RQ1)

記述統計を基に、MTエンゲージメント質問紙のすべての項目の歪度と尖度が $\pm < 1.96$ (Hair et al., 2019)であることから、データが正規分布していることを確かめた。続いて探索的因子分析(最尤法・オリミン回転)を実施して因子を抽出した。固有値1以上、スクリープロット、平行分析の結果(図1)から、5因子構造を仮定した。表1に各質問項目と因子負荷量を示す。



■図1: MTエンゲージメント因子抽出のスクリープロット

■表1: 質問項目と探索的因子分析の結果

No.	項目	因子					独自性
		1	2	3	4	5	
4	英文を書くことについては、私は機械翻訳のヘビーユーザーである	.87	.05	-.04	.01	.00	.22
2	英文を書くときは機械翻訳を使うことが多い	.85	-.02	.06	.02	-.01	.27
5	機械翻訳は英文を書くのに欠かせない	.78	-.02	-.06	.01	.01	.37
3	自分で英文をうまく書けないと思ったときは、機械翻訳を積極的に使う	.64	.06	.10	.06	-.02	.51
1	和英辞書の代わりに機械翻訳を積極的に使っている	.55	-.09	.04	.09	.00	.68
17	機械翻訳を使うときは、意図した英文が出力されるまで日本語を修正する	.09	.71	-.05	-.03	-.01	.51
15	機械翻訳を使うときは、わかりやすい日本語で入力する	-.02	.63	.05	-.05	-.08	.60
16	機械翻訳を使うときは、日本語の主語を補いながら入力する	.05	.63	.01	-.03	-.09	.60
19	機械翻訳の使い方について、自分なりに試行錯誤してみる	-.04	.54	.10	.20	.13	.53
14	機械翻訳を使うときは、何通りか日本語を入力してみる	.12	.41	.25	-.06	.15	.61
8	機械翻訳で出力された英文に誤りがないかチェックし、必要があれば英語を直す	-.03	-.00	.82	-.02	-.02	.33
6	機械翻訳で出力された英文が正しいかどうか、読み返す	.01	-.01	.78	.02	-.02	.39
7	機械翻訳に入力した日本語と、出力された英語を必ず見比べる	.05	.01	.77	-.01	-.00	.38
22	機械翻訳を使うと英文を書くのが苦痛でなくなる	-.00	-.05	.01	.93	-.01	.16
21	機械翻訳を使うとスラスラ英文が書けるので気分がいい	.16	.07	-.11	.71	-.01	.40
23	機械翻訳を使えるなら、喜んで英文を書く	-.07	.05	.11	.55	.09	.65
11	自分が使っている機械翻訳以外に類似のアプリがあるか、友人や教員、学外の人に聞いてみる	.03	.02	-.01	-.02	.74	.46
10	機械翻訳の使い方について、友人や教員、学外の人に聞いてみる	-.01	.01	-.02	-.03	.70	.52
13	機械翻訳の使い方について、友人や教員と議論をする	-.06	-.03	-.04	.03	.68	.54
12	友人と一緒に機械翻訳を使って英文を書く	.34	-.13	-.02	-.02	.46	.68
9	機械翻訳で出力された英文の中に知らない単語があれば、その意味や用法について、辞書で調べる	-.09	.14	.34	.00	-.01	.83
18	機械翻訳の画面上にあるさまざまな機能を使ってみる	-.06	.36	.07	.20	.25	.67
20	異なる機械翻訳を試して、自分に合うものを探す	-.04	.32	.02	.21	.25	.71
24	機械翻訳の使い方について知りたい	-.15	.13	.03	.35	.20	.77
累積因子寄与率		.13	.22	.31	.40	.49	

注: 因子負荷量は > .40, 太字は因子負荷量 > .50 を表す。

分析の結果から、十分な因子負荷量(>.50)を示さなかった6つの項目(9, 12, 14, 18, 20, 24)は除外した。続いて因子を名づけた。第1因子はMTを積極的かつ頻繁に使う項目が多いことから「行動」とした。方略使用に関する項目は、第2因子の、入力時にMTが訳しやすいように日本語を修正する項目と、第3因子の、MT出力の誤りを修正する項目に分かれたため、前者を「プリエディット認知」、後者を「ポストエディット認知」と名づけた。第4因子はMT使用への態度と感情を表しているため「情意」とした。第5因子はMTについてのやりとりを表しているため「社会」とした。これらの因子は先行研究のエンゲージメントの理論と一致したが、主体性を想定した項目18と20は十分な因子負荷量を得られず、項目19は認知の前方略に分類されたため、主体性に関する因子は抽出されなかった。

3.3.2 MTエンゲージメントの信頼性と妥当性(RQ2)

RQ1で抽出された因子構造の信頼性と妥当性を調べるため、最尤法による検証的因子分析を行った。表2にMTエンゲージメント尺度の因子構造(因子負荷量、残差、信頼性係数、AVE、因子間相関)を示す。項目1, 13, 17, 19, 23は残差が.40より大きい(Hair et al., 2019)、因子負荷量が.50を超えているため残した。これらは項目1の「和英辞書の代わりにMTを積極的に使っている」など、質的データから得られた重要な特徴が含まれているためである(Yuasa & Takeuchi, in press)。その結果、1つの因子を構成するために最小限必要な3つの項目を保持した(Hair et al., 2019)。内的一貫性を検討するクロンバックの α 係数は、5つの因子すべてで.73から.89と閾値である.70を満たした。次にIn'nami and Koizumi(2011)を基にモデルの適合基準を用いて構成的妥当性を検証した。 $\chi^2 = 265.34(df = 125, p < .01)$, RMSEA = .06(90%信頼区間:.05–.07)(< .08), CFI = .95(> .90), TLI = .94(> .90), SRMR = .05(< .08)となり、良好な適合指数(Brown, 2015)が得られた。収束妥当性を示すAVEはプリエディット認知が.41と若干低いが、4

■表2: 検証的因子分析によるMTエンゲージメント尺度の因子構造

因子	項目番号	因子負荷量	残差	α 係数	AVE	因子間相関			
						1	2	3	4
1. 行動	4	.88	.23	.89	.64	-			
	2	.85	.27						
	5	.84	.29						
	3	.77	.41						
	1	.60	.65						
2. プリエディット認知	16	.79	.38	.73	.41	.44***	-		
	15	.76	.43						
	17	.53	.73						
	19	.51	.74						
3. ポストエディット認知	6	.84	.30	.86	.68	.12	.57***	-	
	7	.82	.32						
	8	.81	.34						
4. 情意	22	.81	.34	.75	.51	.42***	.23***	.06	-
	21	.78	.39						
	23	.56	.69						
5. 社会	11	.80	.37	.77	.55	.01	.09	.09	.25***
	10	.76	.43						
	13	.63	.61						

注:*** = $p < .001$. AVE = average variance extracted.

つの因子は期待値である .50を満たした(Hair et al., 2019)。さらに各因子のAVEの平方根を計算したところ、すべての値が因子間相関よりも大きかったことから弁別妥当性が確認された。最後に、併存的妥当性を評価するため、MTエンゲージメント尺度とMT自己効力感尺度($\alpha = .80$)の相関を調べたところ、ピアソン相関係数は $r = .36$ ($p < .001$)と中程度に近く、併存的妥当性は弱く支持された。したがって、MTエンゲージメント尺度の信頼性と妥当性は共に満たされていることが確かめられた。

4 調査2: MTエンゲージメントの量的・質的調査

4.1 目的

調査2では大学生英語学習者がMTを使って英文を書くときのエンゲージメントをマクロとミクロの視点から調査する。説明的順次デザインによる混合法(“explanatory sequential mixed methods design”; Creswell & Creswell, 2023, p. 16)を採用し、質問紙による量的データの分析後に、インタビューによる質的データを加えて解釈をする。数値では見えない個人に焦点をあてることで(Mackey & Gass, 2022)、全体像をより鮮明に記述するためである。量的データの分析にはクラスタ分析を用いる。クラスタ分析とは、あるサンプルを変数の組み合わせが似たグループ(クラスタ)に分類する多変量解析法である(Dörnyei, 2007)。第二言語習得研究では、動機づけや学習方略などの個人差要因によって学習者をグループに分けることで、学習者がたどる異なる学習経路を表すために利用されている(Staples & Biber, 2015)。調査1で作成した5因子のMTエンゲージメント尺度の回答データをクラスタ分析し、各因子の強さにより学習者を分類することで、グループ毎の学習者の特徴を明らかにする。さらにグループから数名ずつインタビューを行うことで、MTエンゲージメントをより具体的に記述する。そこで以下のリサーチクエスチョン(RQ)を設定した。

- RQ1 大学生英語学習者は MTエンゲージメントの5因子(行動、プリエディット認知、ポストエディット認知、情意、社会)の強さによって、どのようなグループに分かれるか。
- RQ2 これらのグループの MT自己効力感はどのように異なるか。
- RQ3 各グループの学生はどのようにMTを使っているか。

4.2 方法

4.2.1 参加者

参加者は、調査への協力に同意した、調査者が所属する大学の15学部448名の学生(男性234名、女性214名)である。非英語専攻の学生を対象とした必修英語クラスの1年生(192名)と2年生(256名)で、英語習熟度の判定にはGTEC-LRによるプレメントテスト(500点満点)を利用した。テストの結果($N = 434$, $M = 190.30$, $SD = 52.26$)、点数の分布からおおよそ9割がCEFR A2レベルを占め、少数のCEFR A1とB1レベルを含む、平均的な非英語専攻の大学生英語学習者を表していた。

データ収集に際しては、質問紙への回答は無記名とし、データはすべて統計的に処理して個人が特定されないことを説明した。調査への協力に同意した学生はスマートフォンを使ってGoogle Formsに回答した。さらに回答した参加者からインタビューへの協力者を募集したところ、7名の学生が応じた。インタビュー参加者には実施の当日、調査の目的、データの収集方法、個人情報保護や匿名性などを紙面と口頭で説明し、書面による同意を得た。データ収集に先立ち、調査者が所属する大学院の倫理委員会

に許可を得た。

4.2.2 実施手順

調査2は、質問紙による量的調査とインタビューによる質的調査の2つの調査で構成される。量的調査では、2023年11月から12月にかけて、調査1で作成した「MTエンゲージメント」尺度と「MT自己効力感」尺度の2つの質問紙を配布して回答データを得た。質問紙の尺度は分散を広げて差を可視化するため、6段階尺度(1. まったく当てはまらない 2. 当てはまらない 3. あまり当てはまらない 4. やや当てはまる 5. 当てはまる 6. とてもよく当てはまる)を採用した。学年、学部、性別などの基本情報のほかに、MTの使用頻度を6段階(1. まったく使わない 2. ほとんど使わない 3. あまり使わない 4. ときどき使う 5. よく使う 6. いつも使う)で尋ねる項目も設けた。質的調査では2024年1月から2月にかけて、7名の協力者に半構造化インタビューを対面およびZoomで実施した。所要時間は1人およそ30分を予定し、合計180.9分(平均25.8分)の音声データを得た。

4.2.3 分析方法

質問紙回答から得たデータはJASP(ver.0.18)を使い、階層クラスタ分析(ユーグリッド距離・ワード法)により回答者をエンゲージメントの強さによりグループ分けした(RQ1)。その後、分散分析によりグループ間に差があることを確かめた。さらに各グループの自己効力感に違いがあるかを調べるために、分散分析を行った(RQ2)。続いて、学生個人のMT使用を調べる(RQ3)ために、半構造化インタビューで得た音声データを自動文字起こしツール(Notta AI)と手作業を併用してテキスト化した。テキストデータからMTエンゲージメントを表す部分を抽出し、内容に基づいて6つのエンゲージメントコード(行動、プリエディット認知、ポストエディット認知、情意、社会、主体性)を作成し、定義した。定義に基づき、7人から抽出した合計112個のテキストデータの抜粋に6つのコードでラベルづけした。客観的な判断を得るために、第二言語習得分野に属さないもう1人の分析者がラベルづけを行い、2者間の分析の一致度を確かめた。その結果、106個(94.6%)のラベルが一致し、6つの相違項目は話し合いによって合意に至った。

4.3 結果

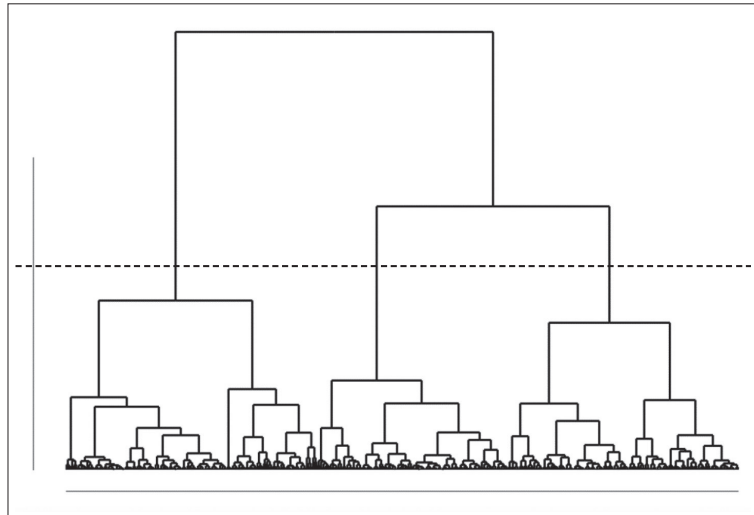
4.3.1 MTエンゲージメントによるグループ分け(RQ1)

データクリーニングにより、テストスコアの欠如や外れ値などを除いた434名のサンプルを使用した。MTの使用頻度を6段階でたずね、その結果、いつも使う($n = 63$)、よく使う($n = 215$)、ときどき使う($n = 132$)の合計が410名で全体の94.5%を占めた。一方、あまり使わない($n = 12$)、ほとんど使わない($n = 8$)、まったく使わない($n = 4$)の合計は24名(5.5%)だった。程度の差はあるが、調査対象となった学生の94.5%がMTを使っているようだ。まったく使わないと回答した学生も、質問紙の回答に1の「まったく当てはまらない」以外の選択肢を含んでいたため、すべてのデータを分析に用いた。表3に質問紙回答から得た各項目の平均と標準偏差を示す。行動とプリエディット・ポストエディットの両認知エンゲージメントは平均値が高かった。情意エンゲージメントは平均値が比較的高い傾向が見られたが、社会エンゲージメントは他の変数に比べると際立って低かった。

■表3: MTエンゲージメント尺度の項目と平均・標準偏差

No.	質問紙項目	M	SD
行動エンゲージメント($\alpha = .89$)			
4	英文を書くことについては、私は機械翻訳のヘビーユーザーである	4.37	1.18
2	英文を書くときは機械翻訳を使うことが多い	4.53	1.03
5	機械翻訳は英文を書くのに欠かせない	4.50	1.10
3	自分で英文をうまく書けないと思ったときは、機械翻訳を積極的に使う	4.85	0.96
1	和英辞書の代わりに機械翻訳を積極的に使っている	4.76	1.10
プリエディット認知エンゲージメント($\alpha = .70$)			
17	機械翻訳を使うときは、意図した英文が出力されるまで日本語を修正する	4.26	1.18
15	機械翻訳を使うときは、わかりやすい日本語で入力する	4.75	1.07
16	機械翻訳を使うときは、日本語の主語を補いながら入力する	4.68	1.06
19	機械翻訳の使い方について、自分なりに試行錯誤してみる	3.99	1.08
ポストエディット認知エンゲージメント($\alpha = .84$)			
8	機械翻訳で出力された英文に誤りがないかチェックし、必要があれば英語を直す	4.77	1.05
6	機械翻訳で出力された英文が正しいかどうか、読み返す	4.84	1.06
7	機械翻訳に入力した日本語と、出力された英語を必ず見比べる	4.80	1.02
情意エンゲージメント($\alpha = .83$)			
22	機械翻訳を使うと英文を書くのが苦痛でなくなる	4.01	1.14
21	機械翻訳を使うとスラスラ英文が書けるので気分がいい	3.85	1.16
23	機械翻訳を使えるなら、喜んで英文を書く	3.61	1.33
社会エンゲージメント($\alpha = .81$)			
11	自分が使っている機械翻訳以外に類似のアプリがあるか、友人や教員、学外の人に聞いてみる	2.86	1.32
10	機械翻訳の使い方について、友人や教員、学外の人に聞いてみる	2.63	1.26
13	機械翻訳の使い方について、友人や教員と議論をする	2.52	1.32

続いて5つのMTエンゲージメントの強さによって参加者をグループに分けるため、階層クラスタ分析(ユークリッド距離・ワード法)を行った。デンドログラム(図2)と各グループの人数、解釈が可能かどうかの判断から、3クラスタによる分類を採用した。



■図2: MTエンゲージメント・クラスタのデンドログラム

3つのグループ間の差を調べるための分散分析を行った。表4にグループ毎の各変数の平均と標準偏差, F 値, 効果量を示す。

■表4: 分散分析の結果

変数	グループ1 ($n=136$) $M(SD)$	グループ2 ($n=100$) $M(SD)$	グループ3 ($n=198$) $M(SD)$	F 値	効果量 ¹ (η^2)	チューキーの 多重比較
行動	4.05 (1.00)	4.35 (0.69)	5.10 (0.59)	83.76***	.28	1-3, 2-3
プリエディット認知	3.80 (0.71)	4.42 (0.65)	4.85 (0.61)	103.26***	.32	1-2, 1-3, 2-3
ポストエディット認知	3.95 (0.85)	5.03 (0.64)	5.28 (0.60)	154.12***	.42	1-2, 1-3
情意	3.55 (0.91)	2.96(0.77)	4.47 (0.82)	118.40***	.36	1-2, 1-3, 2-3
社会	3.08 (0.96)	2.01(0.75)	2.72 (1.20)	31.16***	.13	1-2, 2-3

注: *** = $p < .001$ 1効果量: $> .14$ = 大, $.06 \sim .14$ = 中, $< .06$ = 小。

3つのエンゲージメント・グループは, すべての変数に有意差が見られた。効果量については, 行動, 認知(プリエディット・ポストエディット), 情意が大に対して, 社会のみが中程度だった。これらの結果に基づき, 3つのグループの特徴を解釈し, グループ1を「低エンゲージメント」, グループ2を「中エンゲージメント」, グループ3を「高エンゲージメント」と名づけた。表5にグループの特徴を示す。

■表5: MTエンゲージメント・グループの特徴

グループ名	人数 (割合)	MT使用頻度	特徴
1. 低エンゲージメント	$n=136$ (31.3%)	$M = 4.29$ ($SD = 1.01$)	社会以外のエンゲージメントが低い。 特に行動と認知がグループの中で一番低い。
2. 中エンゲージメント	$n=100$ (23.0%)	$M = 4.58$ ($SD = 0.70$)	行動と認知は3グループの中間だが, 情意と 社会がグループの中で一番低い。
3. 高エンゲージメント	$n=198$ (45.6%)	$M = 5.03$ ($SD = 0.74$)	社会以外のすべてのエンゲージメントが グループの中で一番高い。

グループの特徴は、高エンゲージメント・グループがサンプルの半数近くを占め、4つのエンゲージメント変数とMTの使用頻度が一番高かった。このグループは積極的にMTを使っている可能性がある。一方、全体の3分の1弱を占める低エンゲージメント・グループは、行動と認知が一番低く、MT使用頻度も相対的に低いことから、MT使用にはそれほど積極的でない可能性がある。中エンゲージメント・グループは情意と社会が一番低いことから、MTを使っているが、MTに対してあまり前向きな態度ではない可能性がある。

4.3.2 MTエンゲージメントによるグループの自己効力感(RQ2)

3つのMTエンゲージメント・グループの自己効力感に対する差を調べた。表6にMT自己効力感尺度の質問紙項目と、回答から得た各項目の平均と標準偏差を示す。

■表6: MT自己効力感尺度の項目と記述統計

No.	質問紙項目 ($\alpha=.84$)	M	SD
1	機械翻訳を使って英文を書くとき、正しい文法で書くことができると思う	4.01	0.98
2	機械翻訳を使って英文を書くとき、自然な文章を書くことができると思う	3.54	1.05
3	機械翻訳を使って英文を書くとき、読み手に簡潔でわかりやすい英文を書くことができると思う	3.94	1.03
4	機械翻訳を使って英文を書くとき、自分の意見や考えを適切に表現することができると思う	3.79	1.08
5	機械翻訳を使って英文を書くとき、適切な単語を選ぶことができると思う	3.67	1.07

グループ毎の平均は、低エンゲージメント ($M=3.65$, $SD = 0.74$), 中エンゲージメント ($M=3.34$, $SD = 0.74$), 高エンゲージメント ($M=4.12$, $SD = 0.76$) となった。1元配置の分散分析を行った結果、3グループ間の平均の差は有意となった [$F(2, 431) = 39.25$ ($p < .001$, $\eta^2 = .15$)]。チューキー法による多重比較を行ったところ、低エンゲージメントと高エンゲージメント、中エンゲージメントと高エンゲージメントの平均値に有意な違いが見られた。高エンゲージメント・グループのMT自己効力感が一番高く、中・低エンゲージメント・グループは低かった。

4.3.3 インタビューによる学生個人のMT使用(RQ3)

MTエンゲージメントの各グループ内の学生がどのようにMTを使っているのかを、半構造化インタビューにより明らかにした。インタビューでは事前に準備した質問内容(資料1)を引き出すことを念頭におきながら、調査者と初対面の学生がリラックスして話せるよう、自然な対話の流れを心がけた。協力者の7名の学生は、中エンゲージメント・グループに3名、高エンゲージメント・グループに4名が属した。低エンゲージメント・グループに属するインタビュー参加者はいなかった。表7に参加者のプロフィールを示す。

■表7: インタビュー参加者のプロフィール

グループ	名前*	性別	年齢	学年	GTEC L-R**	よく使うMTツール
中エンゲージメント	Aoi	女	19	2	201	Google 翻訳, DeepL
	Mika	女	19	2	226	Google 翻訳
	Saeko	女	20	2	195	Google 翻訳
高エンゲージメント	Koki	男	19	1	79	Google 翻訳
	Naomi	女	19	1	105	Google 翻訳
	Taka	男	19	1	89	Google 翻訳
	Yoko	女	19	1	119	Google 翻訳, Patago

*仮名を使用。 **GTEC L-R スコアは500点満点

参加者は全員、Google翻訳を使っていたが、Aoi と Yoko はそれぞれ別のMTツール(DeepLとPatago)を併用していた。インタビューからより詳細なエンゲージメントが明らかになった。表8～表13にテキストデータの抜粋を要約した。名前の横の数字は会話のやりとりの順番を示す。

行動エンゲージメント(どんなときに、もしくは何のためにMTを使うか)では、英語のライティング課題だけでなく、さまざまな理由でMTを使っていた(表8)。

■表8: 行動エンゲージメント

中エンゲージメント	高エンゲージメント
・寮の留学生と交流するとき (Aoi, 54) ・バイト先で外国人客の対応に、覚えているフレーズで対処できないとき (157)。仏語の教科書から使える動詞・名詞を探すため (Mika, 161) ・オンライン授業について行くため (Saeko, 100)	・ラグビーの国際試合で海外の選手と交流するとき (Koki, 75) ・英語日記を書くときにより良い表現を探すため (Naomi, 198) ・英作の課題をするとき (Taka, 116) ・英作文の課題でうまく表現できないとき (Yoko, 110)

Aoi, Mika, Koki は外国人とコミュニケーションを図るためにMTを使っていた。Aoi は寮での留学生との会話にMTを使っていた。以下にインタビューの抜粋を示す。()内は調査者が補足した。

抜粋 1

機械翻訳を使わずに自分の英語力で話してみようって思うこともあるんですけど、会話が途切れたり…相手も「え？」って戸惑っているときは、もうすぐに機械翻訳を使っています (Aoi, 140)。

Saeko は1年次に、コロナ禍以降も続いたネイティブ講師によるオンライン授業について行くため、必死にMTを使っていたことを語った(抜粋2)。

抜粋 2

(大学に)入学したとたん、(授業で)英語で書いたり話したりしようって。コミュニケーションの練習をしたことなかったのに。もう(パソコンの)横で翻訳しながらしゃべらないと追いつけないし(100)。そこからどんどん使うようになって便利さもわかってしまって (Saeko, 102)。

7人の中でMika, Naomi, Yoko は自分で書いた英文を改善するためにMTを使っていた。英語で「うまく表現できないときのモヤモヤ感」(27)が嫌いなMikaは、理由をこう説明した(抜粋3)。

抜粋 3

自分の中にない表現を使うときに、やっぱり調べないと(英語の)知識は増えないから(79)。受験のとき以外は、多分、何を学ぶかより、どうやって調べるかのほうが大事だと思います (Mika, 174)。

プリエディット認知エンゲージメントからは、全員が何通りかの日本語を入力し直していることがわかった(表9)。

■表9: プリエディット認知エンゲージメント

中エンゲージメント	高エンゲージメント
・書き言葉で入力する (Aoi, 88) ・日本語の表現を色々変えてみる (Mika, 109) ・わかりやすい日本語で入力する (151)。使いたい構文を考えて日本語を入れる (Saeko, 134)	・意図した英語が出るまで日本語を入力する (Koki, 133) ・教科書のような日本語で入力する (Naomi, 167) ・日本語の語順(副詞の位置など)を変える (Taka, 142) ・シンプルで形式的な日本語で入力する (Yoko, 137)

特に Aoi, Mika, Saeko, Naomi, Yoko の5名は、簡潔でわかりやすい日本語で入力する、口語ではなく書き言葉で入力するなど、MT が出力しやすい日本語 (translation-friendly writing) を意識していた。

ポストエディット認知エンゲージメントから、参加者全員がMTの英文を読んで確認していることがわかった(表10)。

■表10: ポストエディット認知エンゲージメント

中エンゲージメント	高エンゲージメント
- 日本語を Google 翻訳に入力し、その出力を DeepL に入れ、日本語が適切か確かめる (Aoi, 72) - 知らない語は知っている語に置き換えたり (97)、ウェブで調べる (105)。英語の受動・能動態を変える (101)。逆翻訳で意図した意味か確かめる (Mika, 123) - 意図した構文が出ない、英文が長すぎる (134)。ニュアンスが違うとわかったら、日本語を修正する (Saeko, 137, 153)	- 合っているかわからないけど、時々時制を変える (Koki, 202) - 逆翻訳でニュアンスを確かめる (115)。単語をウェブで調べたり置き換える (121)。文法を直す (Naomi, 161) - いつも一応、自分で英文を確認してから使う (Taka, 121) - Patago の訳出を Google 翻訳に入れ、日本語が合っているか確かめる (86)。知らない英単語は日本語に訳したり (115) ウェブで調べる (Yoko, 125)

参加者は全員、MTの出力は不完全だという認識を持っており、意図した英文になっているか、逆翻訳機能を使ってMTの出力を日本語に戻して内容を確認していた。AoiとYoko はさらに別のMTツールを使って意味が通るかどうかを確認していた。Aoi は寮でのカジュアルな会話には Google 翻訳の方が良いと言いつつ、DeepL を併用している理由を説明した(抜粋4)。

抜粋4

逆翻訳も使いますが、Google だけだったら、DeepL で確かめたときにちょっと言い方がきつくなっていたことがあって(74)。… 英語でも丁寧な言い方とかあると思うので、まず丁寧な日本語で入れるようにして、ほんとうにそれで丁寧な言い方になっているのかは、やっぱり (Google から DeepL に) コピペで確かめます (Aoi, 88)。

文法の修正については、時制を変える (Koki) から、受動態・能動態など文の構造を変える (Mika) まで、ポストエディットには程度に差が見られた。インタビューでは Saeko と Naomi が方略の説明に「ニュアンス」という言葉を用いて、言語形式だけでなく意味にも注意を払っている様子が伺えた。Mika は、MT の英文は必ず自分で手を加えてポストエディットする理由をこう説明した(抜粋5)。

抜粋5

(MTの英文を)そのまま使うと、それは自分の意見なのかって言われたら、ちょっと微妙だなんて思います。ちょっと罪悪感もあって (Mika, 141)。

情意エンゲージメントについては、5人の参加者がMT使用への前向きな態度を示した(表11)。

■表11: 情意エンゲージメント

中エンゲージメント	高エンゲージメント
- Google 翻訳はカメラ機能もあってすごく使いやすい (Aoi, 78) - MT がなかったら単位も取れなかったと思う (Saeko, 127)	- MT はめちゃくちゃいいと思う (Koki, 171) - MT があつたら長い文で日記が書ける (Naomi, 195) - これからも MT に手助けしてもらおう (Taka, 184)

一方、高校時代に電子辞書を使って熱心に英語学習に取り組んでいたSaekoは、MTを使うことへの複雑な心情を表した(抜粋6)。

抜粋6

(オンライン授業で)緊張すると、中学生くらいの文も出てこないし、単語も忘れてしまう。(MTで)早く調べなきゃみたいになって。それで英語を身につけようっていうより、単位を取ることに必死になってしまって(107)。… 余計に(MTを)使うようになってしまいました(Saeko, 109)。

社会エンゲージメントについてはAoi, Taka, Yokoが友達からMTアプリを教えてもらっていた(表12)。特にAoiは、2つのアプリを使って自身の英語が留学生との会話で適切かどうか確かめる方法を寮の日本人の友達から教わっていた。実際に友達と一緒にMTを使う参加者はなく、MTは個人での使用に限られていた。

■表12: 社会エンゲージメント

中エンゲージメント	高エンゲージメント
・寮の友達から Google が出力した英文を DeepL で確かめることを習った(Aoi, 120)	・友達に Google を勧められた(Taka, 91) ・友達に Patago を勧められた(Yoko, 48)

主体性エンゲージメントが見られたのは、Naomi だけだった(表13)。

■表13: 主体性エンゲージメント

高エンゲージメント
英語で日記を書くとき、難しい文や知らない単語を Google 翻訳で調べてノートに書き留める(Naomi, 41)

Naomi は半年前から英語で日記をつけるようになった経緯をこう語った(抜粋7)。

抜粋7

去年の夏休みに大学の短期留学に参加したとき、同じセミナーにいた子がすごく英語が上手だったので、その子がしていることを参考に(英語日記を)始めました(29)。… 自分で(英語で)書けるけど、機械翻訳がなかったら、中学校で習った文法ばかりとか、簡単な文になってしまうと思います(Naomi, 193)。

インタビューの最後に、参加者に質問はないか尋ねたところ、高校時代は熱心に電子辞書を使っていたというMikaは、MTの使用について、こう投げかけた(抜粋8)。

抜粋8

今はもう電子辞書よりも機械翻訳がどんどん主流になっていくので、多分、辞書を持たない子のほうが将来、多くなってくると思うんですけど、そういう時に、学校って機械翻訳とどう付き合っていくのかなって思いますね(Mika, 184)。

このように、インタビューからは参加者がそれぞれの置かれた状況の中で、必要に応じて工夫しながらMTを使っていることが明らかになった。抜粋3と8のように、学習者もMTを使う経験を重ねながら、MTに対する自分なりのビリーフを確立させていくのかもしれない。

5 考察

5.1 MTエンゲージメント尺度の作成(調査1)

探索的因子分析ではMTエンゲージメントを構成する行動、プリエディット認知、ポストエディット認知、情意、社会の5つの因子が抽出された(RQ1)。これらの行動、認知、情意、社会エンゲージメント因子は先行研究による理論と一致した。主体性は、Yuasa & Takeuchi (in press)の質的データで示唆されたため尺度に取り入れたが、因子としては抽出されなかった。主体性エンゲージメントとは、学習意欲が高まる環境を作るための学習者の積極的な言動を指す(Reeve & Shin, 2020)。大学の英語教師の多くは学生のMT使用を支持していないこと(Yamada et al., 2021)が、学生のMT使用への積極的・主体的な行動を妨げているのかもしれない。そのことが、主体性因子が抽出されなかった要因になった可能性がある。主体性エンゲージメントの概念はまだ発展途上である(Patall, 2024)。構成概念を検証するためにはさらなる研究が必要である(Sinatra et al., 2015)。

MTエンゲージメント尺度の妥当性の検証では、尺度の信頼性と妥当性が満たされたが(RQ2)、MTエンゲージメントとMT自己効力感との相関はほぼ中程度で、併存的妥当性は弱い支持だった。これはMTの精度に対する学生のビリーフなど、他の要因が働いたからかもしれない。たとえばKoltovskaia(2020)は、人工知能(Artificial Intelligence; AI)を利用したライティングのフィードバック・ツール(Grammarly)への学習者のエンゲージメントを調べた。その結果、学習者がツールに依存するあまり、誤りを修正する際の方略使用が制限されたり、逆にツールへの懐疑心から方略の使用が増えたことを報告している。MTにおいても同様に、MTの出力精度への認識が、プリエディットやポストエディット方略の使用に影響した可能性がある。今回の調査で扱った要因以外の変数についても、今後、さらに調査すべきだろう。

5.2 MTエンゲージメントの量的・質的調査(調査2)

調査2では、CEFR A2レベルに相当するサンプルの大学生英語学習者をMTエンゲージメント尺度の変数(行動・プリエディット認知・ポストエディット認知・情意・社会)の類似グループに分類した。その結果、低エンゲージメント、中エンゲージメント、高エンゲージメントの3つのグループに分類された(RQ1)。グループ間の自己効力感の違いを比べたところ、高エンゲージメント・グループが高く、中・低エンゲージメント・グループは低かった(RQ2)。質問紙項目についての分散分析の結果(表4)とインタビューの結果から(RQ3)、学生は教師が想像する以上にMTの英文を確かめたり修正したりしていた。これはHellmich and Vinall (2023)の、フランス語やスペイン語を学ぶ米国の大学生を対象にした調査結果と一致している。特に全体の半数近く(45.6%)を占める高エンゲージメント・グループはMTを肯定的に捉え、より積極的に使っている可能性がある。このことはサウジアラビアや韓国、インドネシアなどのEFL学習者を対象とした先行研究でも報告されている(Almusharraf & Bailey, 2023; Briggs, 2018; Murtisari et al., 2019)。インタビューからは参加者が方略を使ってMTの英文を修正していることが伺えた。方略使用の程度や意識には差があるようだが、この結果は方略使用と自己効力感(Graham et al., 2020; Teng & Zhang, 2020)、エンゲージメントと自己効力感(Cai & Xing, 2023; Tsao, 2021)に正の相関があることと一致する。

中エンゲージメント・グループは認知エンゲージメントが比較的高く、参加者は方略使用について明確に語っていた。一方、Saekoがオンライン授業で孤立する中、MTを使っていたこと(抜粋2)や、MTに頼ることで英語が身につかないかもしれないという恐れ(抜粋6)、MikaのMT使用に対する「罪悪感」という言葉(抜粋5)は、このグループの情意と社会エンゲージメントの低さを説明している。その要因には、MTを使うことへの倫理的呵責(“ethical qualms”; Klekovkina & Denié-Higney, 2022, p.107)が影響し

ているのかもしれない。

5.3 教育的示唆

インタビューの参加者は全員、先行研究と同様に、MTの出力が不完全であることを認識しながらも (Vinall et al., 2023), 自分で書く英文よりも良い (Murtisari et al., 2024) からと課題にMTを使っていた。“MT-enabled L2 writing” (MTを活用したL2ライティング; Jiang et al., 2024, p. 1) が標準になる時代には、習熟度が低い学習者にとっては、成果物の正確さはもはやL2学習の成果とはみなされないだろう (Klekovkina & Denié-Higney, 2022)。MTの助けを借りながら、最終的なプロダクトにどのようにたどり着いたか、プロセスがより重要である。そこに教師による介入の余地がある。Lee (2021) はCEFR A2レベルの韓国人大学生に1学期間、MTを使ったL2ライティングを指導した。その結果、教師やクラスメイトの足場かけにより、メタ言語知識/言語への気づき (metalinguistic awareness) が向上し、L2ライティングへの自信が増した。

今回の質問紙調査では、社会エンゲージメントの低さが際立った。学習者同士で一緒にMTを使うというインタビュー参加者もいなかった。このことは、教師がMTの使用を認めていないため、MTをオープンに使うことへの倫理的呵責が影響しているのかもしれない。インタビューでは、Koki (202) の「合っているかわからないけど、時々時制を変える」や、Taka (121) の「一応、自分で英文を確認するが、それが正しいかどうかはわからない」などの表現 (表10) に、自身でMTの正確さを判断する自信のなさが表れていた。CEFR A2レベルの学習者は、教師の足場かけの元、協働でMTの英文を修正しながら適切な使い方を学ぶことが必要ではないか。この体験がMTを使って正しい英文が書けるという自己効力感を高め、さらにエンゲージメントを高める (Schunk & DiBenedetto, 2022)。学習者によるMTを使った言語使用についてのやりとりは、Svalberg (2009) が提唱する「言語へのエンゲージメント」 (“engagement with language”) を促進する可能性がある。Svalberg (2018) は「言語へのエンゲージメント」を言語への気づきが高まるプロセスと呼んでいる。言語への気づきはL2学習に不可欠である (Schmidt, 1990)。

Vinall et al. (2023) は、教師と学習者が対話しながらMT使用の方針を学習目的に合わせて柔軟に決めることを提案している。MTの不完全さは、教師と学習者がオープンに議論することで、学びの機会を創出することができる。習熟度が低い学習者にとっては、不完全なL2でやりとりをするよりも、MTを使ってより通じる英文へと磨き上げる活動は、有意義なL2学習になると確信している。自分がほんとうに伝えたい内容について、L1とL2を比較し、語彙や言語形式について議論することは、学習者の知的好奇心を刺激し、エンゲージメントを高めるだろう。

6 今後の課題

今回の調査には限界点もある。MTエンゲージメント尺度の作成では、男女比や学部の違いに配慮したが、一つの大学のみからの参加者を対象にしたので、バイアスがかかる可能性は否めない。他大学の学生も対象とすることで、妥当性を高めることができるだろう。因子として抽出されなかった主体性エンゲージメントに関しては、定義と項目の見直しも必要ではないかと考える。

インタビュー調査では、全体の3割弱を占める低エンゲージメント・グループに属する協力者がいなかったため、このグループの詳細を明らかにすることはできなかった。また7名の参加者からインタビューによりMTの使い方を聞き出したが、実際に使用している場面を確認したわけではない。より詳細に学生個人のMTの使い方を理解するためには、刺激着想法などの研究手法を併用することが望ましい (Hellmich & Vinall, 2023; Yuasa & Takeuchi, 2024)。

7 まとめ

本調査は、CEFR A2レベルに相当する大学生英語学習者のMT使用の現状を、エンゲージメントの枠組みに基づいてマクロとミクロの視点から明らかにした。質問紙調査では、MTエンゲージメント尺度の5つの因子(行動、プリエディット認知、ポストエディット認知、情意、社会)の強さにより、参加者は低・中・高のエンゲージメント・グループに分類された。中エンゲージメントと高エンゲージメント・グループを合わせたおよそ7割の参加者は、プリエディットやポストエディットなどの方略を用いてMTを使っていた。エンゲージメントが高いグループは自己効力感も高く、エンゲージメントが低いグループは自己効力感も低かった。すべてのグループで社会エンゲージメントが際立って低いことから、MTは個人での使用に限られているようだ。インタビューから、参加者がそれぞれの目的に合わせて、工夫しながらMTを使っていることがわかった。一方、MTの英文の正確さを十分に判断できていない可能性や、MTを使うことへの倫理的呵責も見えた。習熟度が低い学習者は、教師や学習者同士のやりとりを通して、MTを使いながら正しい英文へと磨いていく体験が必要である。このプロセスはL2学習に必要な言語への気づきを高めることができる。そこで初めて、MTは学習ツールとして利用することができるだろう。本調査がAI時代にふさわしい教授法を模索するきっかけになればと願う。

謝辞

本調査を実施するにあたり、貴重な機会とご支援をくださいました公益財団法人 日本英語検定協会と関係者の皆さま、選考委員の皆さまに心よりお礼申し上げます。特に研究助言者の寺内一先生には貴重なご助言をいただき、感謝申し上げます。また、今回の質問紙調査は、関西大学の多くの学生と先生方のご協力によって実現できました。さらにインタビューでは、7名の学生に機械翻訳についての貴重なご意見をいただきました。心より感謝申し上げます。なお、本報告書に関して、開示すべき利益相反関連事項はありません。

引用文献

- Almusharraf, A., & Bailey, D. (2023). Machine translation in language acquisition: A study on EFL students' perceptions and practices in Saudi Arabia and South Korea. *Journal of Computer Assisted Learning*, 39(6), 1988-2003. <https://doi.org/10.1111/jcal.12857>
- Bai, B., & Wang, J. (2023). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, 27(1), 207-228. <https://doi.org/https://doi.org/10.1177/1362168820933190>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front Public Health*, 6. <https://doi.org/10.3389/fpubh.2018.00149>
- Bowker, L., & Buitrago, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, 14(1), 3-24. <https://doi.org/10.29140/jaltcall.v14n1.221>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Cai, Y., & Xing, K. (2023). Examining the mediation of engagement between self-efficacy and language achievement. *Journal of Multilingual and Multicultural Development*, 1-13. <https://doi.org/10.1080/01434632.2023.2217801>
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. <https://doi.org/10.1016/j.system.2020.102408>
- Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, 100(S1)(Suppl. 1), 64-80. <https://doi.org/10.1111/modl.12302>
- Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *The Journal of AsiaTEFL*, 17(1), 182-

引用文献

193. <https://doi.org/10.18823/asiatefl.2020.17.1.11.182>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage Publications.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals*, 51(4), 779-795. <https://doi.org/10.1111/flan.12366>
- Gally, T. (2019). The implications of machine translation for English education in Japan. *JACETSIG-ELE Journal* 6(2), 1-14. <https://ci.nii.ac.jp/naid/40021993334/>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning and Technology*, 26(2), 5-24. <https://doi.org/10.1257/73474>
- Goetze, J., & Driver, M. (2022). Is learning really just believing? A meta-analysis of self-efficacy and achievement in SLA. *Studies in Second Language Learning and Teaching*, 12(2), 233-259. <https://doi.org/10.14746/ssllt.2022.12.2.4>
- Graham, S. (2022). Self-efficacy and language learning: What it is and what it isn't. *The Language Learning Journal*, 50(2), 186-207. <https://doi.org/10.1080/09571736.2022.2045679>
- Graham, S., Woore, R., Porter, A., Courtney, L., & Savory, C. (2020). Navigating the challenges of L2 reading: Self-efficacy, self-regulatory reading strategies, and learner profiles. *The Modern Language Journal*, 104(4), 693-714. <https://doi.org/10.1111/modl.12670>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hellmich, E. A., & Vinall, K. (2023). Student use and instructor beliefs: Machine translation in language education. *Language Learning and Technology*, 27(1), 1-27. <https://hdl.handle.net/10125/73525>
- Hiver, P., Al-Hoorie, A. H., & Mercer, S. (2021). *Student engagement in the language classroom*. Multilingual Matters.
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 28(1), 201-230. <https://doi.org/10.1177/13621688211001289>
- In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly*, 8(3), 250-276. <https://doi.org/10.1080/15434303.2011.582203>
- Jiang, L., Yu, R., & Zhao, Y. (2024). Theoretical perspectives and factors influencing machine translation use in L2 writing: A scoping review. *Journal of Second Language Writing*, 64, 101099. <https://doi.org/10.1016/j.jslw.2024.101099>
- Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *The Modern Language Journal*, 108(2), 515-533. <https://doi.org/10.1111/modl.12924>
- Klekovkina, V., & Denié-Higney, L. (2022). Machine translation: Friend or foe in the language classroom? *L2 Journal*, 14(1), 105-135. <https://doi.org/10.5070/I214151723>
- Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, 28(1), 663-682. <https://doi.org/10.1007/s10639-022-11194-2>
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google translate in academic writing courses? *The EUROCALL Review*, 26(2), 50-57. <https://doi.org/10.4995/eurocall.2018.10140>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175. <https://doi.org/10.1080/09588221.2018.1553186>
- Lee, Y.-J. (2021). Still taboo? Using machine translation for low-level EFL writers. *ELT Journal*, 75(4), 432-441. <https://doi.org/10.1093/elt/ccab018>
- Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: an international journal*, 46(3), 517-528. <https://doi.org/10.2224/sbp.7054>
- Liu, M., Noordin, N., Ismail, L., & Abdrahim, N. A. (2023). Relationship between student engagement and academic achievement in college English education for non-English majors in China. *International Journal of Learning, Teaching and Educational Research*, 22(8), 203-232. <https://doi.org/10.26803/ijlter.22.8.12>
- Lo, S. (2023). Neural machine translation in EFL classrooms: Learners' vocabulary improvement, immediate vocabulary retention and delayed vocabulary retention. *Computer Assisted Language Learning*, 1-20. <https://doi.org/10.1080/09588221.2023.2207603>
- Mackey, A., & Gass, S. M. (2022). *Second language research: Methodology and design* (3rd ed.). Routledge.
- Mercer, S. (2019). Language learner engagement: Setting the scene. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 643-660). Springer.
- Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.
- Murtisari, E. T., Kristianto, A. K., & Bonar, G. (2024). Self-directed use of machine translation among language learners: Does it lead to disruptive L2 avoidance? *Foreign Language Annals*, 1-24. <https://doi.org/10.1111/flan.12768>
- Murtisari, E. T., Widiningrum, R., Branata, J., & Susanto, R. D. (2019). Google Translate in language learning: Indonesian EFL

引用文献

- students' attitudes. *The Journal of Asia TEFL*, 16(3), 978-986. <https://doi.org/10.18823/asiatefl.2019.16.3.14.978>
- Negishi, M., Takada, T., & Tono, Y. (2013). A progress report on the development of the CEFR-J. In E. D. Galaczi & C. J. Weir (Eds.), *Exploring language frameworks: Proceedings of the ALTE Kraków conference, July 2011* (pp. 135-163). Cambridge University Press.
- O'Neill, E. M. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL-EJ*, 20(1), 154-177. <https://old.callej.org/journal/20-1/O'Neill2019.pdf>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578. <https://doi.org/10.3102/0034654306004543>
- Patall, E. A. (2024). Agentic engagement: Transcending passive motivation. *Motivation Science*, 10(3), 222-233. <https://doi.org/10.1037/mot0000332>
- Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice*, 59(2), 150-161. <https://doi.org/10.1080/00405841.2019.1702451>
- Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4), 257-267. <https://doi.org/10.1016/j.cedpsych.2011.05.002>
- Reschly, A. L., & Christenson, S. (2022). *Handbook of research on student engagement* (2nd ed.). Springer.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schunk, D. H., & DiBenedetto, M. K. (2022). Self-efficacy and engaged learners. In A. L. Reschly & S. Christenson (Eds.), *Handbook of research on student engagement* (2nd ed., pp. 155-170). Springer.
- Shin, D., & Chon, Y. V. (2023). Second language learners' post-editing strategies for machine translation errors. *Language Learning and Technology*, 27(1), 1-25. <https://hdl.handle.net/10125/73523>
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, 50(1), 1-13. <https://doi.org/10.1080/00461520.2014.1002924>
- Staples, S., & Biber, D. (2015). Cluster analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 243-274). Routledge.
- Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, 56, 18-34. <https://doi.org/10.1016/j.esp.2019.07.001>
- Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, 90, 102221. <https://doi.org/10.1016/j.system.2020.102221>
- Svalberg, A. M. L. (2009). Engagement with language: Interrogating a construct. *Language Awareness*, 18(3-4), 242-258. <https://doi.org/10.1080/09658410903197264>
- Svalberg, A. M. L. (2018). Researching language engagement: Current trends and future directions. *Language Awareness*, 27(1-2), 21-39. <https://doi.org/10.1080/09658416.2017.1406490>
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48, 100701. <https://doi.org/10.1016/j.jslw.2019.100701>
- Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5-6), 510-526. <https://doi.org/10.1080/09588221.2018.1527361>
- Tsao, J.-J. (2021). Effects of EFL learners' L2 writing self-efficacy on engagement with written corrective feedback. *The Asia-Pacific Education Researcher*, 30(6), 575-584. <https://doi.org/10.1007/s40299-021-00591-9>
- Vinall, K., Wen, W., & Hellmich, E. A. (2023). Investigating L2 writers, uses of machine translation and other online tools. *Foreign Language Annals*, 57(2), 499-526. <https://doi.org/10.1111/flan.12733>
- Wang, C., & Sun, T. (2020). Relationship between self-efficacy and language proficiency: A meta-analysis. *System*, 95, 102366. <https://doi.org/10.1016/j.system.2020.102366>
- Williams, M., & Burden, R. L. (1997). *Psychology for language teachers: A social constructivist approach*. Cambridge University Press.
- Yamada, M., Langliz, H., Oda, T., Morita, T., Tamura, H., Hiraoka, Y., & Irie, T. (2021). A preliminary survey on the use of machine translation in English education at Japanese universities. *Invitation to Interpreting and Translation Studies*, 23, 139-156. https://doi.org/http://honyakukenyu.sakura.ne.jp/shotai_vol23/No_23_007-Yamada_et-al.pdf
- Yuasa, M., & Takeuchi, O. (2022). Strategic use of Google Translate by Japanese university students of English [Conference presentation]. *Situating Strategy Use: Strategic Learning in an uncertain world (SSU4)*, Wellington, New Zealand.
- Yuasa, M., & Takeuchi, O. (2024). Stimulated recall interviews using digital technology in L2 research. In K. Sadeghi (Ed.), *Handbook of technological advances in researching language learning*. Routledge.
- Yuasa, M., & Takeuchi, O. (in press). Strategic use of machine translation: A case study of Japanese EFL university students.
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). Cambridge University Press.

資料1：半構造化インタビューの質問項目

Part 1: 英語学習について

1. 英語学習は好きですか？
2. その理由は？
3. あなたの習熟度はどれくらいだと思いますか？
4. 英検やTOEICなどの英語の資格試験を受験したことはありますか？

Part 2: ライティング活動について

1. 英語のライティングは好きですか？
2. 高校の英語の授業でライティングの指導は受けましたか？
3. どんな内容でどのくらいの指導がありましたか？
4. ライティングのどこが難しいと思いますか？

Part 3: MTを使う背景や目的, MTの精度について

1. どのMT(アプリ)を使っていますか？
2. いつ頃からMTを使い始めましたか？
3. MTを使うようになったきっかけは何だと思いますか？
4. MTを使う理由や目的は何ですか？
5. MTのどんなところが良いと思いますか？
6. MTのどんなところが良くないと思いますか？
8. MTの訳は正確だと思いますか？

Part 4: MTの使い方について

1. MTを使うとき, 初めから日本語を入力しますか?自分で英文を書いてから使いますか？
2. 入力する日本語を直しますか？
3. MTが出力した英文を自分で見直していますか？
4. MTが出力した英文を自分で修正しますか？
5. 修正するときに, MT以外に辞書などを参照していますか？
6. どんなもの(辞書・ウェブサイトなど)を参照しますか？
7. MTの使い方を誰かに教えてもらったことはありますか？
8. 今後もMTを使い続けたいですか？
9. ChatGPTなどの, 他のAIツールを使ったことがありますか？